A Mean Field Game Approach for Distributed Control of Thermostatic Loads Acting in Simultaneous Energy-Frequency Response Markets

Antonio De Paola, *Member, IEEE,* Vincenzo Trovato, *Member, IEEE,* David Angeli, *Fellow, IEEE*
Goran Strbac, *Member, IEEE*

***Abstract*—This paper proposes a novel distributed solution for the operation of large populations of thermostatically controlled loads (TCLs) providing frequency support. A game-theory framework is adopted, modelling the TCLs as price-responsive rational agents that schedule their energy consumption and allocate frequency response provision in order to minimize their operational costs. The novelty of this work lies in the use of mean field games to abstract the complex interactions of large numbers of TCLs with the grid and in the introduction of an innovative market structure, envisioning distinct price signals for electricity and response. Differently from previous approaches, such prices are not designed ad hoc but are derived instead from an underlying system scheduling model.**

***Index Terms*—Game theory, distributed control, power system control, load management, thermal energy storage.**

NOMENCLATURE

A. Indices

τ	Index of generation technology
i	Time index of integration grid
j	Temperature index of integration grid
k	Initial temperature index of integration grid

B. Parameters

$c_{1,\tau}$	No-load cost of generation technology τ (£/MWh)
$c_{2,\tau}$	Linear production cost of generation technology τ (£/MWh)
$c_{3,\tau}$	Quadratic production cost of generation technology τ (£/MW^{2}h)
r_τ	Headroom for FR of generation technology τ
s_τ	FR/dispatch level slope of generation technology τ
$G_\tau^{\max}$	Installed capacity of generation technology τ (MW)
ΔG_L	Loss of generation in frequency event (MW)
D	FR damping parameter (s)
$\Delta f_{\text{qss}}^{\max}$	Quasi-steady-state frequency constraint (Hz)
Δf_{rcf}	Frequency parameter in RoCoF constraint (Hz)
$\Delta f_{\max}$	Maximum frequency deviation (Hz)
t_{rcf}	Time parameter in RoCoF constraint (s)
f_0	Nominal frequency (Hz)
h_τ	Inertia constant of generation technology τ (s)
h_L	Inertia of the infeed generation loss (s)
σ	Thermal time constant of the TCLs (s)
T_{off}	Ambient temperature (°C)
ζ	Heat exchange parameter of the TCLs (°C/W)
P_{ON}	Power consumption of TCL in ON state (W)
T_{MIN}	Minimum feasible temperature of TCL (°C)
T_{MAX}	Maximum feasible temperature of TCL (°C)
t_{FIN}	Duration of considered time horizon (h)
N	Number of considered TCLs
U_0	Aggregate steady-state power of TCLs (MW)
α	Discomfort cost parameter (£/h(°C)^2)
$\bar{T}$	Target temperature of discomfort cost (°C)
β	Anti-synchronization cost parameter (£/h(°C)^2)
Δt	Time integration step (s)
ΔT	Temperature integration step (°C)

C. Functions

$p(\cdot)$	Electricity price (£/Wh)
$\rho(t)$	Availability fee awarded for FR provision (£/Wh)
$U_{\text{TCL}}(\cdot)$	Aggregate power consumption of TCLs (MW)
$R_{\text{TCL}}(\cdot)$	Aggregate FR allocated by TCLs (MW)
$H(\cdot)$	Online generation vector
$G(\cdot)$	Generated power vector (MW)
$R(\cdot)$	Vector of frequency response from generators (MW)
$\bar{U}(\cdot)$	Inflexible demand profile (MW)
$\varphi(\cdot,\cdot)$	Minimized total generation costs per unit of time (£/h)
$\hat{R}(\cdot)$	Aggregate system FR (MW)
$\hat{H}(\cdot)$	System post-fault inertial response (MW $\cdot$ s^2)
$T(\cdot)$	Internal temperature of TCL (°C)
$I(\cdot)$	Temperature variation of TCL (°C)
$u(\cdot)$	Power consumption of TCL (W)
$\mathfrak{f}(\cdot,\cdot)$	Temperature dynamics of TCL (°C/s)
$\lambda(\cdot)$	Fraction of power consumption allocated as FR
$r(\cdot,\cdot)$	Allocated FR of TCL (W)
$J(\cdot)$	Cost function of TCL (£)
$\bar{I}(\cdot)$	Target temperature variation (°C)
$\Psi(\cdot)$	Terminal cost function (£)
$V(\cdot,\cdot,\cdot)$	Value function for cost-minimization of TCL (£)
$u^*(\cdot,\cdot,\cdot)$	Optimal feedback control of TCL (W)
$m(\cdot,\cdot,\cdot)$	State distribution of TCLs population

I. INTRODUCTION

In the context of flexible demand-side participation, thermostatically controlled loads (TCLs) such as refrigerators or air conditioners have been consistently investigated in the last

A. De Paola (corresponding author - e-mail: a.de.paola@bath.ac.uk) is with the University of Bath, V. Trovato is with EDF Energy R&D UK Centre, D. Angeli and G. Strbac are with Imperial College London. V. Trovato is also with Imperial College London and D. Angeli is also with the University of Florence. This work was supported by the Leverhulme Trust under Grant [ECF-2016-394].

few years. The intrinsic flexibility of TCLs has been exploited to provide several ancillary services and support system operation. For instance, TCLs may perform energy arbitrage [1] or provide frequency response [2], mitigating severe frequency deviations due to rare events (e.g. sudden generators outages). The paper focuses on these two aspects and neglects, for the time being, other potential TCLs applications such as balancing mechanisms for frequency regulation [3].

A. Relevant Work

A popular approach to exploit TCL flexibility in power system operation envisages a cooperative coordination of the appliances. Under distributed control implementations, previous works [4], [5] have demonstrated the collective ability for a populations of TCLs to contain post-fault frequency deviations. The TCL response in [4] highly depended on system conditions and it suffered from unintended synchronization and uncontrolled steady-state recovery. Solutions to these issues eventually reduced the TCL contribution [5]. Other works ([1], [6], [7]) described the TCL population by means of aggregate models, e.g. a battery-like representation in [6]. The success of nicely matching aggregate power profiles with system dynamics was counterbalanced by the need for a real-time, centralized and (often) bi-directional control to determine single TCL operation. The stochastic controller in [8] offered an effective compromise between the advantages of distributed control and system-oriented goal-setting capability.

The aforementioned cooperative approaches exhibit some common features: i) they do not ensure an optimal policy for each TCL, which could be better off by optimizing its own comfort/cost; ii) they require an intermediate entity (e.g. aggregators) to translate the TCLs' flexibility into financial rewards, iii) they do not fully preserve privacy, as users might have to divulge their preferences and behavioural patterns.

For all these reasons, competitive schemes have been investigated for control and optimization of the devices' operation. Within this framework, each load responds to pricing/control signals by selecting the operation strategy that achieves the best trade-off between costs and comfort of its user.

A market-based approach is considered in [9]–[11], where flexible loads respond to market prices that arise from unit commitment/generation scheduling and implicitly account for the impact of demand response on system operation. A wide range of solutions, including load saturation [12] and price corrective mechanisms [13], has been proposed to facilitate market clearing. Novel pricing mechanisms that account for investment costs [14] and ad-hoc bidding strategies for flexible demand integration [15] have also been evaluated.

In parallel, another research stream has investigated game-theoretical approaches that consider simplified price structures and instead focus on the competitive interactions between the loads. The seminal papers [16], [17] have proposed iterative updates of broadcast price/control signals to achieve a desirable coordination of the flexible loads, characterized as a game equilibrium. These approaches have been recently expanded in multiple directions, considering for example increased decentralization [18] or reduced iterations [19].

B. Contributions

The objective of this paper is to bridge the gap between the aforementioned research directions on competitive methods for flexible demand coordination (market and game-based). In this respect, we propose a novel price-based framework that combines a detailed modelling of system scheduling and market prices with a rigorous game-theoretical formulation.

The main contribution of the paper is the application of mean field games (MFG), a novel analytical tool that is used to describe the interactions between large populations of TCLs. These are modelled as rational agents pursuing the minimization of their individual operating cost. By approximating the size of the appliances population as infinite and neglecting the impact of the single small load (while still accounting for the effect of the whole population) this approach provides a scalable and compact characterization for the equilibrium solution of the TCLs coordination problem. At equilibrium, the TCLs are applying their best response to specific price signals and, by doing so, they are inducing those very same prices in the system.

Mean field games have been previously applied in power system contexts, considering for example electric vehicles [20], storage [21] or TCLs [22]. However, to the best of our knowledge, this is the first paper that implements some fundamental elements in a mean field game framework:

- Formulation of a multi-objective perspective in which the agents not only exchange energy with the system but also provide ancillary services, such as frequency response (FR). The trade-off between these two (potentially conflicting) objectives is determined in a purely economic framework: each TCLs will aim to minimize the difference between the energy cost and the revenues for FR provision.
- The price signals used for the coordination of the appliances are derived from a unit commitment (UC) model, rather than designed ad hoc or obtained through simplified closed-form expressions [16]–[19]. As a result, the considered prices accurately reflect the cost of electricity and the value of response, accounting for different generation technologies and detailed FR requirements. As the broadcast prices are equal for all TCLs, fairness of the proposed scheme is ensured.
- Novel anti-synchronization techniques are introduced to preserve diversity of the agents' population and facilitate convergence to equilibrium through the introduction of auxiliary state variables and the design of ad hoc cost functions.

The paper also discusses the distributed methods (based on iterative price updates) that can be used to converge to the desired equilibrium solution and presents detailed schemes for the numerical resolution of the TCLs mean field game equations. The performance of the proposed approach is then evaluated in simulation, quantifying the potential value of the proposed control scheme for the individual agents.

C. Paper Structure

The rest of the paper is structured as follows: Section II describes relevant elements and assumptions of the UC model and the TCL dynamics. Section III contains the MFG

formulation, whereas Section IV discusses schemes and algorithms for the numerical resolution of the coordination problem. Section V discusses the practical implementation of the proposed strategy and Section VI presents a case study on a large-scale power system, comparing the results of the proposed approach with alternative paradigms. Finally, Section VII contains conclusive remarks.

II. Modelling

The proposed modelling framework is summarized in Fig. 1. A population of TCLs is coordinated through the broadcast of two signals p and ρ, which represent respectively the price of electricity and the availability fee awarded for FR provision. In turn, the behaviour of the TCLs is summarized by their aggregate power consumption U_{TCL} and their total frequency response R_{TCL}. These quantities modify the UC and therefore impact the prices p and ρ.

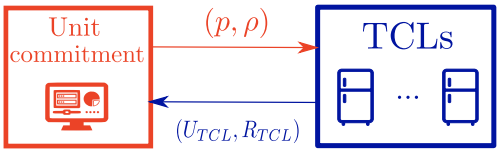


Fig. 1. Diagram of the interactions between TCLs and unit commitment.

A. Unit Commitment

The UC determines generation scheduling decisions (in terms of energy and FR) in order to minimize the short-term operating costs of the system. The result of the UC can be interpreted as the solution of a traditional centralized market mechanism for energy and FR, if one assumes inelastic demand and perfect competition [12]. Under this market framework, the prices of energy and FR broadcast to the TCLs (p and ρ, respectively) correspond to the relevant Lagrange multipliers of the UC optimization problem.

For simplicity, a Linear Programming (LP) formulation is adopted for the UC problem. Considering a generic generation technology τ and assuming that the size of single plants included in τ is quite smaller than the aggregate installed capacity of τ, commitment decisions can be extended to the fleet and expressed by continuous variables $H_\tau(t) \in [0,1]$. Besides a significant reduction in computational time granted by LP UC rather than binary UC and mixed integer linear programming (MILP) UC, authors in [23] demonstrated that a LP model captures, with relatively high precision, all the relevant system-level scheduling requirements, namely generation-demand balancing and simultaneous allocation of frequency-related control to deal with fault events.

For consistency with the mean field game formulation presented in Section III, the proposed deterministic formulation of the UC problem is time-continuous. To simplify the analysis, inter-temporal constraints (e.g. generation ramping) are neglected, allowing to solve the dispatch problem on a instant-by-instant basis. We wish to emphasize that the distributed control strategy for coordination of TCLs presented in this paper can be easily extended to more complex models of UC.

Hence, let N_τ denote the number of generation technologies available in the system. We indicate by $H_\tau(t)$, $G_\tau(t)$ and $R_\tau(t)$ the on line generation (as fraction of the total installed and available capacity), the generated power and the allocated response of technology τ at time $t \in [0, t_{\mathrm{FIN}}]$. The associated commitment, generation and response vectors are denoted by $H(t) = [H_1(t), \ldots, H_{N_\tau}(t)]$, $G(t) = [G_1(t), \ldots, G_{N_\tau}(t)]$ and $R(t) = [R_1(t), \ldots, R_{N_\tau}(t)]$, respectively. In the present analysis, the allocated response corresponds to the contribution of each technology to system frequency requirements, i.e. the allocated power headroom that can be used in case of frequency events.

If the total power consumption and allocated response of the TCLs at time t are equal to $U_{\mathrm{TCL}}(t)$ and $R_{\mathrm{TCL}}(t)$, respectively, the UC at time t can be described through the following optimization problem:

$$\varphi(U_{\mathrm{TCL}}(t), R_{\mathrm{TCL}}(t)) = \min_{H(t),G(t),R(t)} \sum_{\tau=1}^{N_\tau} c_{1,\tau} \cdot H_\tau(t) G_\tau^{\max} + c_{2,\tau} \cdot G_\tau(t) + c_{3,\tau} \cdot G_\tau^2(t) \quad (1)$$

with $c_{1,\tau}$ [£/MWh] as no-load cost, $c_{2,\tau}$ [£/MWh] and $c_{3,\tau}$ [£/MW^2h] as production costs. The objective function (1) corresponds to total generation costs per unit of time and it is subject to the constraints below:

$$\sum_{\tau=1}^{N_\tau} G_\tau(t) - \bar{U}(t) - U_{\mathrm{TCL}}(t) = 0 \quad (2a)$$

$$0 \le H_\tau(t) \le 1 \quad (2b)$$

$$R_\tau(t) - r_\tau \cdot H_\tau(t) \cdot G_\tau^{\max}(t) \le 0 \quad (2c)$$

$$R_\tau(t) - s_\tau \cdot [H_\tau(t) \cdot G_\tau^{\max}(t) - G_\tau(t)] \le 0 \quad (2d)$$

$$\Delta G_{\mathrm{L}} - D[\bar{U}(t) + U_{\mathrm{TCL}}(t) - R_{\mathrm{TCL}}(t)] \Delta f_{\mathrm{qss}}^{\max} - \hat{R}(t) \le 0 \quad (2e)$$

$$2\Delta G_{\mathrm{L}} \cdot t_{\mathrm{rcf}} \cdot t_{\mathrm{d}} - t_{\mathrm{rcf}}^2 \cdot \hat{R}(t) - 4\Delta f_{\mathrm{rcf}} \cdot t_{\mathrm{d}} \cdot \hat{H}(t) \le 0 \quad (2f)$$

$$\bar{q}(t) - \hat{H}(t) \cdot \hat{R}(t) \le 0 \quad (2g)$$

$$\mu \cdot r_\tau \cdot H_\tau(t) \cdot G_\tau^{\max}(t) - G_\tau(t) \le 0 \quad (2h)$$

where (2a) equals generation and aggregated demand (i.e. the system inelastic demand $\bar{U}(t)$ and the TCL flexible demand $U_{\mathrm{TCL}}(t)$). The fundamental assumption for LP UC previously discussed is confirmed by (2b). The amount of response allocated by each generation technology is limited by the headroom $r_\tau \cdot H_\tau(t) \cdot G_\tau^{\max}(t)$ in (2c) and the slope s_τ linking the FR with the dispatch level (2d).

Similarly to [24], the constraints (2e)-(2g) guarantee secure frequency deviations following the sudden generation loss ΔG_L. In particular, (2e) allocates enough FR (with delivery time t_d) such that the quasi-steady-state frequency deviation is above $\Delta f_{\mathrm{qss}}^{\max}$, with D accounting for the damping effect introduced by load [24]. As for GB practice on the rate of change of frequency (RoCoF) limit, it is imposed in (2f) that at t_{rcf} the frequency deviation is above Δf_{rcf}. For a more compact notation, we have introduced the aggregate system response $\hat{R}$ and the system post-fault inertial response (IR) $\hat{H}$, defined as follows:

$$\hat{R}(t) = \sum_{\tau=1}^{N_\tau} R_\tau(t) + R_{\mathrm{TCL}}(t) \quad (2i)$$

$$\hat{H}(t) = \sum_{\tau=1}^{N_\tau} \frac{h_\tau \cdot H_\tau(t) \cdot G_\tau^{\max} - h_{\mathrm{L}} \Delta G_{\mathrm{L}}}{f_0}. \tag{2j}$$

Relevant quantities in the above expressions include the nominal frequency f_0, the inertia constant h_τ of technology τ and the inertia h_L of the infeed generation loss (no longer supporting the system). Finally (2g) constraints the maximum tolerable frequency deviation Δf_{nad}, following the formulation and methodology presented in [24] and [25]. For a given system condition at time t (e.g. $\bar{U}(t)+U_{\mathrm{TCL}}(t)$, t_{d}, ΔG_{L}), there exists a unique $\bar{q} = \hat{R} \cdot \hat{H}$ such that $|\Delta f_{\mathrm{nadir}}| = \Delta f_{\max}$. Note that (2g) is a convex function in the range of possible $\hat{R}$ and $\hat{H}$. Hence, we linearize it around $n = 10$ points $x_k = (\hat{R}^k, \hat{H}^k)$, chosen such that $g(\hat{R}^k, \hat{H}^k) = \bar{q}$. At low demand conditions (typically during night hours), (2e)-(2g) would require high levels of IR and FR. Constraint (2h) prevents trivial unrealistic solutions that may arise in the proposed LP formulation. In particular, it avoids high values of committed generation $H(t)$ in correspondence with low (even zero) generation dispatch $G(t)$. Such configuration would provide large amounts of inertial response (proportional to $H(t)$) and primary response $R(t)$, but it would most likely violate the threshold for minimum stable generation of the considered plants. To avoid this undesired outcome, the generation dispatch $G(t)$ must be μ times higher than the maximum available FR.

Having characterized the minimized operation cost φ in (1) as a function of the current power consumption $U_{\mathrm{TCL}}(t)$ and allocated response $R_{\mathrm{TCL}}(t)$ of the TCL population, it is now possible to introduce the following quantities:

$$p(t) = \left.\frac{\partial \varphi(U,R)}{\partial U}\right|_{\substack{U=U_{\mathrm{TCL}}(t) \\ R=R_{\mathrm{TCL}}(t)}} \rho(t) = \left.\frac{\partial \varphi(U,R)}{\partial R}\right|_{\substack{U=U_{\mathrm{TCL}}(t) \\ R=R_{\mathrm{TCL}}(t)}} \tag{3}$$

As mentioned above, the result of the UC can be interpreted as a market solution. Within this framework, $p(t)$ in (3) represents the marginal cost of accommodating one additional unit of demand, which in this specific case corresponds to an additional unit of power consumption from the TCLs. As a result, $p(t)$ can be considered as the price of electricity paid by the TCLs at time t. With a similar reasoning, $\rho(t)$ in (3) represents the marginal saving of allocating one less unit of FR, assuming that such response is being provided by the TCLs instead. It follows that $\rho(t)$ can be interpreted as the availability price paid to the TCLs for FR provision. The prices p and ρ in (3) can be expressed as Lagrange multipliers of the associated constraints in (2) or, equivalently, can be calculated through a variational sensitivity analysis.

B. Thermostatically controlled loads

A standard first-order dynamical model is used to describe the time-evolution of the internal temperature T of a TCL:

$$\dot{T}(t) = \mathfrak{f}(T(t), u(t)) = -\frac{1}{\sigma}\left(T(t) - T_{OFF} + \zeta u(t)\right) \tag{4a}$$

$$T(0) = \tilde{T} \tag{4b}$$

where $u(t)$ represents the instantaneous power consumption of the TCL (either equal to 0 or P_{ON}), σ is its thermal time constant, T_{OFF} is the ambient temperature, ζ incorporates the physical model of heat exchange (with $T_{OFF} - \zeta u(t)$ indicating the asymptotic cooling temperature) and $\tilde{T}$ is the initial TCL temperature. In addition to T, an additional state variable I is considered, with $I(t) = T(t) - \tilde{T}$ denoting the variation of internal TCL temperature in the time interval $[0,t]$. The dynamic evolution of I is straightforward to derive:

$$\dot{I}(t) = \dot{T}(t) \tag{5a}$$

$$I(0) = 0. \tag{5b}$$

The auxiliary state I ensures that multiple TCLs with equal temperature $T(t)$ at time t remain distinguishable by means of distinct values of $I(t)$, as long as their initial temperature $\tilde{T}$ is different. This is crucial to avoid synchronization in the distributed coordination scheme presented in Section III.

At each time $t \in [0, t_{\mathrm{FIN}}]$, the single TCL can either be ON ($u(t) = P_{\mathrm{ON}}$) or OFF ($u(t) = 0$). The maximum amount of FR r that it can allocate as a result corresponds to:

$$r(T(t), u(t)) = \lambda(T(t)) \cdot u(t) \tag{6}$$

where $\lambda : [T_{\mathrm{MIN}}, T_{\mathrm{MAX}}] \to [0,1]$ is a decreasing function of the internal temperature T, with $\lambda(T_{\mathrm{MAX}}) = 0$ and $\lambda(T_{\mathrm{MIN}}) = 1$. For simplicity, in the current analysis a linear expression is considered for λ:

$$\lambda(T) = \frac{T - T_{\mathrm{MAX}}}{T_{\mathrm{MIN}} - T_{\mathrm{MAX}}}. \tag{7}$$

The TCL can only provide response if it is ON ($u(t) = P_{\mathrm{ON}}$) and therefore able to reduce its power consumption in case of a frequency event. The weighting factor $\lambda(T)$ accounts for temperature constraints and ensures that a potential FR provision (with the TCL being in the OFF state for a certain time) does not lead to undesirable temperature values.

Remark 1: In the proposed modelling framework, generation failures and, subsequently, provision of FR by the TCLs are assumed to be extremely rare. For this reason, the variation in cost and temperature dynamics associated to such events are not accounted for in the considered model of normal operation. This means that all TCLs will always allocate their maximum available frequency response r, since there is no advantage in providing less response.

The single TCL is modelled as a self-interested agent that schedules its power consumption $u(\cdot)$ over a certain time horizon $[0, t_{\mathrm{FIN}}]$ on the basis of the signals $p(\cdot)$ and $\rho(\cdot)$ in (3), i.e. the current prices of electricity and response availability. The cost function J minimized by a TCL is expressed as:

$$J(u(\cdot)) = \int_0^{t_{\mathrm{FIN}}} p(t) \cdot u(t) - \rho(t) \cdot r(T(t), u(t)) + \alpha(T(t) - \bar{T})^2 + \beta(I(t) - \bar{I}(t))^2 \, dt + \Psi(I(t_{\mathrm{FIN}})), \tag{8}$$

subject to:

$$T(0) = \tilde{T} \qquad I(0) = 0 \tag{9}$$

$$T(t) \in [T_{\mathrm{MIN}}, T_{\mathrm{MAX}}] \qquad \forall t \in [0, t_{\mathrm{FIN}}] \tag{10}$$

where T_{MIN} and T_{MAX} are the operational temperature bounds of the considered TCL. Each term in (8) is described in detail:

- $p(t) \cdot u(t)$ (*electricity cost*): the instantaneous cost associated to the power consumption $u(t)$.

- $\rho(t)\cdot r(T(t),u(t))$ (*response revenue*): availability fee awarded for FR provision. It corresponds to the product of the availability response price $\rho(t)$ and the allocated response $r(T(t),u(t))=\lambda(T(t))\cdot u(t)$.
- $\alpha(T(t)-\bar{T})^2$ (*T tracking cost*): discomfort term penalizing temperature deviations from some comfort target $\bar{T}$.
- $\beta(I(t)-\bar{I}(t))^2$ (*I tracking cost*): anti-synchronization term that quadratically penalizes deviations of the state variable I from some target signal $\bar{I}$. It is introduced to diversify the behaviour of groups of TCLs that have converged to the same temperature $T(t)$ but have different initial temperatures $T(0)=\tilde{T}$ and therefore heterogeneous values of the state variable $I(t)=T(t)-\tilde{T}$.
- $\Psi(I(t_{\text{FIN}}))$ (*terminal cost function*): imposes periodic constraints, ensuring that the final temperature $T(t_{\text{FIN}})$ of each TCL is equal to its initial value $\tilde{T}$. For example, recalling that $I(t)=T(t)-\tilde{T}$, one can choose $\Psi(I)=\Lambda\cdot I^2$.

We wish to emphasize the multi-objective perspective of the cost function J: each TCL will not only consider the impact of its operational strategy u on its energy costs, but it will also account for the resulting revenues derived from FR allocation and the discomfort caused by mismatches between the internal temperature T and its target value $\bar{T}$.

III. Mean-field game formulation

The TCLs are modelled as individual players that aim at minimizing their cost $J(u)$ in (8) and compete between each other through the changes in prices p and ρ induced by their operational strategies, i.e. their power consumption $u(t)$. It is assumed that the single device is sufficiently small not to significantly impact the system prices. At the same time, the non-negligible effect of the aggregate strategy of the TCL population is taken into account. The behaviour of the TCLs can then be summarized by two coupled partial differential equations (PDEs) that describe the optimal strategy to be followed by the devices and their state evolution over time. To introduce the first equation, the value function V is considered:

$$V(t,T,I)=\min_{u(\cdot)}\int_t^{t_{\text{FIN}}} p(s)\cdot u(s)-\rho(s)\cdot r(T(s),u(s))+ \alpha(T(s)-\bar{T})^2+\beta(I(s)-\bar{I}(s))^2\,ds+\Psi(I(t_{\text{FIN}})), \quad (11)$$

subject to (10) and:

$$T(t)=T \qquad I(t)=I. \quad (12)$$

Note that V, corresponding to the minimized cost sustained on the time interval $[t,t_{\text{FIN}}]$ by a device with state as in (12), can be characterized as the solution of an Hamilton-Jacobi-Bellman (HJB) equation. Denoting by V_x the partial derivative of V with respect to the variable x, it holds:

$$-V_t(t,T,I)=\min_{u\in\{0,P_{\text{ON}}\}}\big[p(t)u-\rho(t)r(T,u)+\alpha(T-\bar{T})^2 +\beta(I-\bar{I}(t))^2+(V_T(t,T,I)+V_I(t,T,I))\mathfrak{f}(T,u)\big] \quad (13a)$$

$$V(t_{\text{FIN}},T,I)=\Psi(I). \quad (13b)$$

Following the dynamic programming principle [26], the value function V is obtained by an integration backward in time, selecting at each time t the control u that minimizes the right-hand side of (13a). The boundary condition (13b) straightly follows from (11) evaluated at $t=t_{\text{FIN}}$. As a result, it is possible to derive the optimal feedback strategy of the TCLs for a certain set of price signals $(p(\cdot),\rho(\cdot))$, defined as a function $u^*(t,T,I)$ of time instant t and current state (T,I):

$$u^*(t,T,I)=\operatorname*{arg\,min}_{u\in\{0,P_{\text{ON}}\}}\big[p(t)u-\rho(t)r(T,u)+\alpha(T-\bar{T})^2 +\beta(I-\bar{I}(t))^2+(V_T(t,T,I)+V_I(t,T,I))\mathfrak{f}(T,u)\big] \quad (14)$$

The second PDE is now introduced to characterize the dynamic evolution of the TCLs population when each device acts rationally and applies the optimal feedback strategy u^*. To this end, let $m(t,T,I)$ denote the state distribution of the devices, with $\int_{I_1}^{I_2}\int_{T_1}^{T_2} m(t,T,I)\,dT\,dI$ corresponding to the fraction of devices with temperature $T(t)\in[T_1,T_2]$ and temperature variation $I(t)\in[I_1,I_2]$ at time t. The distribution m fulfils the following transport equation:

$$m_t(t,T,I)=-[m(t,T,I)\mathfrak{f}(t,u^*(t,T,I))]_T -[m(t,T,I)\mathfrak{f}(T,u^*(t,T,I))]_I \quad (15a)$$

$$m(0,T,I)=m_0(T)\cdot\delta(I) \quad (15b)$$

where δ denotes the Dirac delta function and m_0 represents the initial temperature distribution of the TCLs. Denoted by N the total number of devices, it is straightforward to characterize U_{TCL} and R_{TCL} as the following weighted integrals:

$$U_{\text{TCL}}(t)=N\int_{I_{\text{MIN}}}^{I_{\text{MAX}}}\int_{T_{\text{MIN}}}^{T_{\text{MAX}}} u^*(t,E,I)m(t,E,I)\,dT\,dI \quad (16a)$$

$$R_{\text{TCL}}(t)=N\int_{I_{\text{MIN}}}^{I_{\text{MAX}}}\int_{T_{\text{MIN}}}^{T_{\text{MAX}}} r(T,u^*(t,T,I))m(t,T,I)\,dT\,dI. \quad (16b)$$

The solution of the distributed coordination problem for large populations of TCLs corresponds to the quantities $(p,\rho,V,u^*,m,U_{\text{TCL}},R_{\text{TCL}})$ that satisfy the set of equations (3), (13) - (16). These can be interpreted as a market equilibrium: the population of TCLs autonomously apply the optimal policy u^* in response to the broadcast price signals p and ρ. By doing so, the associated total power consumption U_{TCL} and allocated response R_{TCL} induce the very same prices of energy and response that had been considered in the first place.

IV. Numerical Resolution

A closed-form solution for the coupled PDEs (13) and (15) may not exist and it is in general difficult to calculate. For this reason, the equilibrium of the TCLs coordination problem is calculated through numerical resolution of (1) - (3) and (13) - (16), as discussed in the rest of this section.

A. Integration Schemes

The mean field game equations (13)-(15) are solved over a discretized grid of time and temperature values. To avoid numerical integration issues for the grid points at the border of the feasibility region, a change of state variables is introduced, replacing the pair (T,I) with $(T,\tilde{T})$. In this equivalent representation, the initial temperature $\tilde{T}$ is considered a constant

state component, with $\dot{\tilde{T}}(t)=0$. The original state variable $I(t)$ can always be derived from the new state as $I(t)=T(t)-\tilde{T}(t)$. Denoted by Δt and ΔT the chosen time and temperature discretization steps, the single grid points are described as:

$$(t_i, T_j, \tilde{T}_k) = (i\Delta t, T_{\text{MIN}} + j\Delta T, T_{\text{MIN}} + k\Delta T). \tag{17}$$

The indexes $i = 0,1,\dots,t_{\text{FIN}}/\Delta t$, $j = 0,1,\dots,(T_{\text{MAX}} - T_{\text{MIN}})/\Delta T$ and $k = 0,1,\dots,(T_{\text{MAX}} - T_{\text{MIN}})/\Delta T$ are used to denote the points in the grid that correspond to times t in the considered time horizon $[0, t_{\text{FIN}}]$ and temperature values T and $\tilde{T}$ in the feasible interval $[T_{\text{MIN}}, T_{\text{MAX}}]$. For compactness of notation, any arbitrary function f evaluated at a certain grid point $(t_i, T_j, \tilde{T}_k)$ is equivalently denoted as $f^i_{j,k}$.

The schemes adopted for the numerical integration of the PDEs are now presented. For the HJB equation (13) we have:

$$V^{i-1}_{j,k} = V^i_{j,k} + \min_{u \in \mathscr{U}_j} \left(L^i_{j,k}(u) \right) \Delta t \tag{18}$$

where $\mathscr{U}_j \subseteq \{0, P_{\text{ON}}\}$ denotes the feasible controls at a certain temperature T_j and L in the argument of the min function corresponds to the right-hand side of (13) evaluated on the discretized grid when a certain control u is applied. For the function L, the following expressions are considered:

$$\begin{aligned} L^i_{j,k}(0) = &-\rho^i \cdot r(T_j, 0) + \alpha (T_j - \bar{T})^2 \\ &+ \beta (T_j - \tilde{T}_k - \bar{I}^i)^2 + \left(\frac{V^i_{j+1,k} - V^i_{j,k}}{\Delta T} \right) \mathfrak{f}(T_j, 0). \end{aligned} \tag{19a}$$

$$\begin{aligned} L^i_{j,k}(P_{\text{ON}}) = &\, p^i P_{\text{ON}} - \rho^i \cdot r(T_j, P_{\text{ON}}) + \alpha (T_j - \bar{T})^2 \\ &+ \beta (T_j - \tilde{T}_k - \bar{I}^i)^2 + \left(\frac{V^i_{j,k} - V^i_{j-1,k}}{\Delta T} \right) \mathfrak{f}(T_j, P_{\text{ON}}). \end{aligned} \tag{19b}$$

An upwind method [27] is applied in (19): the partial derivative V_T is approximated with forward differences when $u = 0$ and $\mathfrak{f}(T_j, u) > 0$ and with backward differences when $u = P_{\text{ON}}$ and $\mathfrak{f}(T_j, u) < 0$. This reduces the integration error and avoids issues at the temperature bounds of the discretized grid. As we are considering the equivalent (constant) state variable $\tilde{T}$ rather than I, numerical calculation of V_I in (13a) is not necessary.

From (14) the optimal feedback control u^* on the discretized time/temperature grid is equal to:

$$u^{*i}_{j,k} = \arg\min_{u \in \mathscr{U}_j} L^i_{j,k}(u). \tag{20}$$

To numerically integrate the transport equation (15), some preliminary quantities are introduced, denoting by $\theta(j,k)$ the temperature value obtained starting from T_j and applying the optimal feedback control $u^*(t, T_j, T_j - \tilde{T}_k)$ for Δt time. From (4a), considering a first-order Taylor approximation, it holds:

$$\theta(j,k) = T_j + \mathfrak{f}(T_j, u^*(t_i, T_j, T_j - \tilde{T}_k))\Delta t. \tag{21}$$

In general, $\theta(j,k)$ will correspond to a temperature value between two adjacent grid points $T_{\underline{\gamma}(j,k)}$ and $T_{\bar{\gamma}(j,k)}$. The indexes $\underline{\gamma}(j,k)$ and $\bar{\gamma}(j,k)$ have the following expressions:

$$\underline{\gamma}(j,k) = \left\lfloor \frac{\theta(j,k) - T_{\text{MIN}}}{\Delta T} \right\rfloor \quad \bar{\gamma}(j,k) = \left\lceil \frac{\theta(j,k) - T_{\text{MIN}}}{\Delta T} \right\rceil \tag{22}$$

where the angular brackets denote floor and ceiling functions. The numerical integration of the transport equation (15) can then be performed as in [21, Section IV], initializing m^i at 0 and iterating the following over j and k:

$$m^i_{\underline{\gamma}(j,k),k} = m^i_{\underline{\gamma}(j,k),k} + m^{i-1}_{j,k} \left[1 - \frac{\theta(j,k) - T_{\underline{\gamma}(j,k)}}{\Delta T} \right] \tag{23a}$$

$$m^i_{\bar{\gamma}(j,k),k} = m^i_{\bar{\gamma}(j,k),k} + m^{i-1}_{j,k} \frac{\theta(j,k) - T_{\underline{\gamma}(j,k)}}{\Delta T} \tag{23b}$$

At each time index i, the updated temperature value $\theta(j,k)$ associated to the grid point $(T_j, \tilde{T}_k)$ and to the optimal control $u^*(t, T_j, T_j - \tilde{T}_k)$ is calculated. In a second step, the fraction of devices $m^{i-1}_{j,k}$ is repartitioned at the new time index i between the grid points $(\underline{\gamma}(j,k),k)$ and $(\bar{\gamma}(j,k),k)$, proportionally to the proximity of $\theta(j,k)$ to $T_{\underline{\gamma}(j,k)}$ and $T_{\bar{\gamma}(j,k)}$. This integration scheme ensures mass conservation of the distribution m and an accurate representation of average temperature variations. Canonical schemes, such as the Lax-Friedrich method with viscosity [28], return comparable results.

B. Resolution Algorithm

The numerical resolution of the system of equations (1) - (3), (13) - (16) is performed with the iterative approach detailed in Algorithm 1, solving each single equation and using the obtained result as a starting point for the next calculations.

Algorithm 1 Distributed price-based coordination of TCLs

1) **Initialization.** Total power consumption and FR of the TCLs are set to some initial values $\tilde{U}_{\text{TCL}}$ and $\tilde{R}_{\text{TCL}}$:

$$\tilde{U}_{\text{TCL}}(t) \leftarrow U_0 \qquad \tilde{R}_{\text{TCL}}(t) \leftarrow 0 \quad \forall t \in [0, t_{\text{FIN}}]. \tag{24}$$

2) **Iterative resolution.** Equations are solved sequentially:
   a) Solve the UC (1)-(2) for $U_{\text{TCL}} = \tilde{U}_{\text{TCL}}$, calculating the associated price signals $\tilde{p}$ and $\tilde{\rho}$ according to (3).
   b) Calculate the solution $\tilde{V}$ and $\tilde{u}^*$ of (13) and (14) when $p = \tilde{p}$ and $\rho = \tilde{\rho}$, using numerical schemes (18)-(20).
   c) Calculate the temperature distribution $\tilde{m}$ according to (15) when $u^* = \tilde{u}^*$, using numerical schemes (21)-(23).
   d) Evaluate updated estimates $\hat{U}_{\text{TCL}}$ and $\hat{R}_{\text{TCL}}$ for total power consumption and response allocation of TCLs according to (16), assuming $u^* = \tilde{u}^*$ and $m = \tilde{m}$.
3) **Convergence check.** Evaluate the error:

$$e = \int_0^{t_{FIN}} \left| \tilde{U}_{\text{TCL}}(t) - \hat{U}_{\text{TCL}}(t) \right| + \left| \tilde{R}_{\text{TCL}}(t) - \hat{R}_{\text{TCL}}(t) \right| dt \tag{25}$$

IF $e > \varepsilon$. Repeat step 2) with updated values:

$$\tilde{U}_{\text{TCL}} \leftarrow \nu \cdot \tilde{U}_{\text{TCL}} + (1-\nu) \cdot \hat{U}_{\text{TCL}} \tag{26a}$$

$$\tilde{R}_{\text{TCL}} \leftarrow \nu \cdot \tilde{R}_{\text{TCL}} + (1-\nu) \cdot \hat{R}_{\text{TCL}} \tag{26b}$$

ELSE return the (approximate) equilibrium solution:

$$u^* = \tilde{u} \quad V = \tilde{V} \quad m = \tilde{m} \quad p = \tilde{p} \quad \rho = \tilde{\rho}. \tag{27}$$

The UC equations (1) - (3) in step 2.a) are solved at discrete time instants $t = k\Delta\tilde{t}$. The time step $\Delta\tilde{t}$ is generally greater or equal than Δt used in the integration of (13) - (16), with $\Delta\tilde{t} = \eta\Delta t$ for some $\eta \in \mathbb{N}$. To consistently combine the two different time steps, the price signals $p(t)$ and $\rho(t)$ calculated at each $t = k\Delta\tilde{t}$ are assumed to be constant over the associated η time samples with step Δt. Conversely, the signals U_{TCL} and R_{TCL} are averaged over η time samples with step Δt in order to obtain consistent values when $\Delta\tilde{t}$ is considered instead.

C. Computational Complexity and Scalability

The operations required for the integration schemes presented in (18)-(23) are proportional to the number of grid points defined in (17). As a result, the computational complexity for the numerical integration of the PDEs increases linearly with respect to a reduction in the time step Δt and quadratically when the temperature step ΔT is reduced (as the latter refers to both state variables T and I - or equivalently T and $\tilde{T}$). The same considerations apply to Algorithm 1, as this includes iterated applications of the integration schemes (18)-(23) and of the UC problem (1)-(2), whose complexity is linear with respect to reductions of the associated time step $\Delta\tilde{t}$ and it is not affected by the temperature step ΔT.

We wish to underline the scalability of the proposed resolution method, whose complexity does not explicitly depend on the number N of considered flexible TCLs. This is ensured by the chosen mean field game framework, which allows to characterize the overall behaviour of the devices' population with only two quantities: the optimal feedback control u^* and the state distribution m. The penetration level of the TCLs will only affect indirectly the convergence speed of the algorithm. For higher values of U_0, the impact of U_{TCL} and R_{TCL} on the solution of the UC problem (1)-(2) will be more relevant. As a result, higher values of ν might be necessary in the update equation (26), potentially leading to a slower convergence to equilibrium.

V. Practical Implementation

This section deals with the possibility of deploying the proposed control strategy in a real-world context, considering in particular the implementation schemes and the technical elements required to enable the participation of flexible TCLs to FR.

A. Implementation schemes

The price-based coordination of the TCLs can be implemented in practice through a bi-directional communication scheme between the TCLs and some central coordinator (e.g. the system operator). One effective coordination procedure, represented in Fig. 2, could strictly follow the individual steps of Algorithm 1. In this case, the central coordinator would initially consider some estimate of $\tilde{U}_{\mathrm{TCL}}$ and $\tilde{R}_{\mathrm{TCL}}$, which characterize the aggregate behaviour of the TCLs population (Step 1). On the basis of these quantities, it would solve the UC problem and broadcast to the TCLs the resulting price signals $\tilde{p}$ and $\tilde{\rho}$ (Step 2.a). In turn, the TCLs respond to the

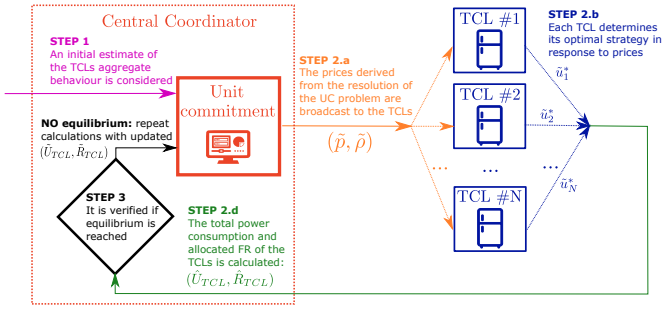


Fig. 2. Iterative implementation scheme for TCLs coordination.

prices by determining their optimal power scheduling $\tilde{u}^*$ (Step 2.b), which they communicate to the central coordinator. In this practical implementation, Step 2.c of the algorithm can be bypassed and the central entity can directly calculate the new aggregate quantities $\hat{U}_{\mathrm{TCL}}$ and $\hat{R}_{\mathrm{TCL}}$ (Step 2.d). In the final Step 3, the central coordinator establishes whether the equilibrium has been reached or a new iteration of Step 2 needs to be performed with updated quantities.

An alternative one-shot implementation strategy, schematically represented in Fig. 3, can also be considered. In this case,

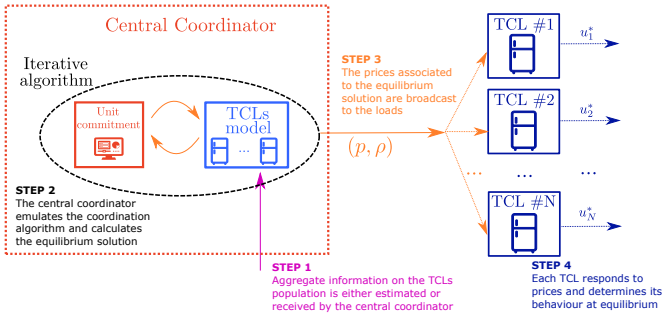


Fig. 3. One-shot implementation scheme for TCLs coordination.

it is assumed (Step 1) that the central coordinator has received (or estimated with sufficient precision), the initial temperature distribution m_0 of the TCLs and the relevant parameters in their dynamics (4a). In other words, the central entity must rely on a database of the most relevant TCLs' models and quantify their penetration in the system. With this information, the central entity is able to internally emulate Algorithm 1 and autonomously calculate the equilibrium solution (Step 2). At this point, it can directly broadcast the resulting price signals to the TCLs (Step 3), which can then calculate their optimal control strategy (Step 4), thus inducing the desired equilibrium solution in one step. It should be mentioned that this one-shot algorithm achieves faster convergence at the price of a reduced customers' privacy since some information on the TCLs (albeit in aggregate form) will have to be divulged.

B. Technical requirements

A high-level diagram of the main technical components and data signals required for participation of flexible TCLs to an integrated energy-FR market is represented in Fig. 4. It can be seen that two main additional components need

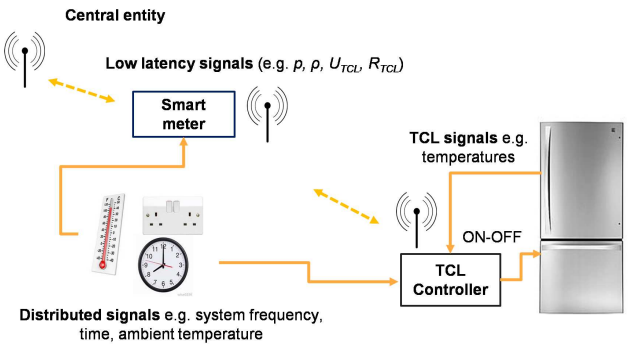


Fig. 4. Technical diagram of TCL with smart controller.

to be considered with respect to traditional configurations: a TCL controller and a smart meter. The TCL controller (by-passing or in parallel to the regular thermostatic controller) performs the calculation of the optimal power schedule on the basis of the received signals and commands the actuator that switches ON and OFF the TCLs. At the moment, these types of controllers are still in a prototype phase but it is thought that, through economies of scale, their price could be as low as 3£ [4]. In the considered context, a smart meter will also be required to record the TCLs' scheduled operation (i.e. profiles of power consumption and FR allocation/provision), communicate it to the central coordinator and calculate the associated energy costs and FR revenues to be claimed. The signals required and exchanged by the different components to operate under the envisioned flexible price-responsive paradigm can be divided in three groups:

1) *TCLs' signals*: these include in particular the TCLs' internal temperature, used to determine the optimal feedback control.
2) *Distributed signals*: (e.g. time, system frequency) required to detect frequency events and associate them to the relevant actions and prices.
3) *Ad-hoc signals*: (e.g. prices, TCLs' schedule) these are exchanged between the central entity and the loads and are required to implement the proposed control strategy.

VI. Simulative Results

The distributed coordination strategy and resolution techniques presented in Section III and IV are tested in simulation, considering a typical 2030 scenario of the GB system as in [24]. The performance of the proposed control strategy is also compared, in terms of operational costs for the system and the TCLs, with a cooperative flexible approach and with the business-as-usual scenario.

A. Case Study

System parameters not yet defined are $\Delta G_{\mathrm{L}} = 1800$ MW, $h_{\mathrm{L}} = 6$ s, $D = 1/50$ s, $f_0 = 50$ Hz, $t_{\mathrm{rcf}} = 0.5$ s, $\Delta f_{\mathrm{rcf}} = 0.5$ Hz, $\Delta f_{\mathrm{nad}} = 0.8$ Hz, $\Delta f_{\mathrm{qss}}^{\max} = 0.5$ Hz and $t_{\mathrm{d}} = 10$. The generation technologies available in the system are nuclear, combined cycle gas turbines (CCGT), open cycle gas turbines (OCGT) and wind. Their characteristics are listed in Table I.

TABLE I
Characteristics of Thermal and Wind Generation

	Nuclear	**CCGT**	**OCGT**	**Wind**
Installed capacity $G_\tau^{\max}$ [MW]	10000	25000	20000	40000
No-load cost $c_{1,\tau}$ [£/MWh]	0	7	8	0
Production cost $c_{2,\tau}$ [£/MWh]	1	25	50	0.5
Production cost $c_{3,\tau}$ [£/MW^{2}h]	0.0005	0.005	0.01	0.0005
Constant of inertia h_τ [s]	6	4	4	-
FR headroom r_τ	-	0.1	0.1	-
FR slope s_τ	-	0.4	0.5	-

For simplicity, in this work the optimization horizon is limited to $t_{\mathrm{FIN}} = 24$ h. Hence, the system inelastic demand $\bar{U}$ reflects a typical winter work-day profile and, together with wind realization (green dotted line in Fig. 7a), is taken from [24]. Further work will extend the optimization horizon (e.g. to one year) to account for the typical daily/seasonal variability of wind and demand when calculating TCL operation and value.

It is assumed that a population of $N = 2 \cdot 10^7$ fridges with built-in freeze compartment operates in the system according to the proposed price-based control scheme. Each device, on the basis of the received price signals p and ρ, schedules its power consumption u on the time interval $[0, t_{\mathrm{FIN}}]$. The initial temperature distribution m_0 in (15b) is uniform and homogeneity of the TCLs is assumed: each fridge has the same parameters (specified in [29]) and the resulting steady-state power consumption of the aggregate TCLs is $U_0 = 790$ MW. Note that the extension of the proposed modelling framework to the case of a finite number of heterogeneous TCL models (each with different parameters) is straightforward to implement through the procedure presented in [21, Section IV].

For the cost function (8) of the individual device, the terminal cost is chosen as $\Psi(I) = \Lambda I^2$ with $\Lambda = 0.001$£/°C^2, penalizing quadratically any temperature variation $I(t_{\mathrm{FIN}}) = T(t_{\mathrm{FIN}}) - T(0)$. The set-point for the state variable I is defined as $\bar{I}(t) = \kappa \sin(\phi t)$, with $\kappa = 1°$ C and $\phi = 1$ h^{-1}. The choice of a sinusoidal function for $\bar{I}$ aims at emulating the normal variational trend in the TCLs internal temperature, which will usually oscillate during the day. In this respect, we wish to emphasize that the sensitivity of the final results with respect to the parameters κ and ϕ is significantly low, as discussed at the end of this section. The remaining cost function parameters are $\alpha = 2 \cdot 10^{-5}$£/h(°C)^2, $\beta = 2 \cdot 10^{-4}$£/h(°C)^2 and $\bar{T} = -17.5$°C.

B. Market Equilibrium Solution

Algorithm 1 has been applied to the presented scenario, choosing $\Delta\tilde{t} = 0.5$ h as time step in the resolution of (1)-(3) and $\Delta t = 7.5$ s and $\Delta T = 0.15°$ C as discretization steps in (18) - (23). With $\nu = 0.95$ and $\varepsilon = 1.5$ MWh in the **Convergence check** phase, the algorithm terminates after 75 iterations. A MATLAB environment has been used for the simulations and the resolution precedure requires about 11 minutes on a standard laptop HP EliteBook with an Intel(R) i5 CPU (2.60 GhZ) and 8 GB of RAM.

The resulting profile of total power consumption $U_{\mathrm{TCL}}^{(n)}$ and total allocated response $R_{\mathrm{TCL}}^{(n)}$, at different iterations n of the resolution algorithm, are reported in Fig. 5. The electricity

prices $p^{(n)}$ and response availability prices $\rho^{(n)}$ (equal to the Lagrange multipliers of the active constraints (2a) and (2g), respectively) are shown in Fig. 6. Note that, after some initial oscillations, convergence is achieved to the equilibrium power profiles $U_{\mathrm{TCL}}^{(75)}$ and $R_{\mathrm{TCL}}^{(75)}$ (Fig. 5) and to the price signals $p^{(75)}$ and $\rho^{(75)}$ (Fig. 6), represented in black. In general, the total power consumption U_{TCL} is higher during the first hours of the day, when the electricity price p is lower. The significant increase of p at about $t = 5$ h is accompanied by a sharp reduction of the TCL power consumption at similar times. The power U_{TCL} then oscillates during the day in order to maintain feasible levels of the internal temperature T of the TCLs. It can also be seen that U_{TCL} has a steep decrease starting at about $t = 16$ h, when the price p starts to increase consistently over the 100£/MWh value. The decrease of U_{TCL} continues over the next three-hour interval of high-price values, including the price peak at about $t = 17$ h.

Similarly to U_{TCL}, the aggregate frequency response R_{TCL} allocated by the TCLs will in general be higher during the early hours of the day (up to $t = 5$ h). This will occur because the power/price relationship between the response availability fee ρ and the allocated response R_{TCL} is the opposite of the one between the electricity price p and allocated power consumption U_{TCL}. In broad terms, R_{TCL} will tend to be high when ρ is high and will generally be low when ρ is low. Note that the opposite occurs for U_{TCL}, which is high when the price p is low. This follows from the different sign of the two price terms p and ρ in the cost function J of the single load in (8). In particular, p appears with a plus sign in the cost J because it represents the price paid by the TCLs for power consumption. Conversely, ρ is considered with a minus sign since it denotes the availability fee at which the loads are rewarded for their FR allocation. When a significant price variation occurs at $t = 5$ h (with p increasing and ρ decreasing), both U_{TCL} and R_{TCL} will tend to be lower.

From expression (6), the frequency response r allocated by each TCL is unequivocally determined by its internal temperature T and its power consumption u. This means that, as the TCLs aim at fulfilling their temperature requirements during the day by consuming power when electricity is cheap, they will also automatically provide a certain amount of FR, even at times when the availability fee ρ for response allocation is low, e.g. after $t = 5$ h. Moreover, the broad price/power trends discussed above can be identified only when the differences and variations in prices are significant. In general, the behaviour of the TCLs and the associated aggregate quantities U_{TCL} and R_{TCL} will not only follow the oscillations of the prices p and ρ but they will also be impacted by the other cost terms in (8) and by the temperature bounds presented in (10).

The TCL impact on system commitment decisions and consequent energy/FR dispatch levels is also analyzed and displayed in Fig. 7, comparing the case when TCL flexibility is enabled (solid lines) with a no-flexibility scenario (dashed lines) where $U_{\mathrm{TCL}}(t) = U_0$ and $R_{\mathrm{TCL}}(t) = 0$, $\forall t \in [0, t_{\mathrm{FIN}}]$.

Fig. 7a shows the generated power from each technology. Since the TCLs demand is much smaller than inelastic demand and its variation across the day is limited (see Fig. 5), its

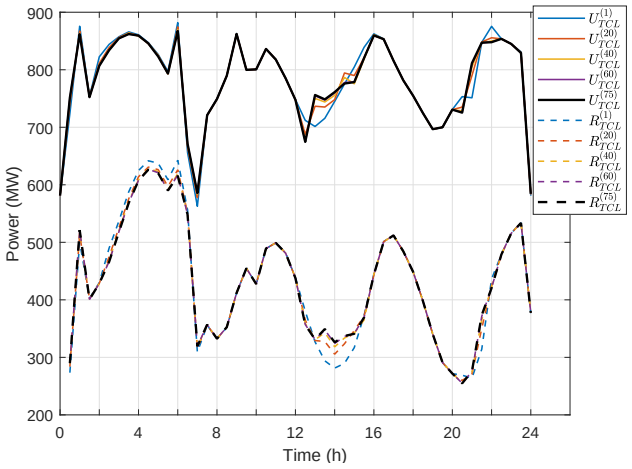


Fig. 5. Total power consumption $U_{\mathrm{TCL}}^{(n)}$ (bold lines) and allocated response $R_{\mathrm{TCL}}^{(n)}$ (dashed lines) at different iterations n of the coordination algorithm.

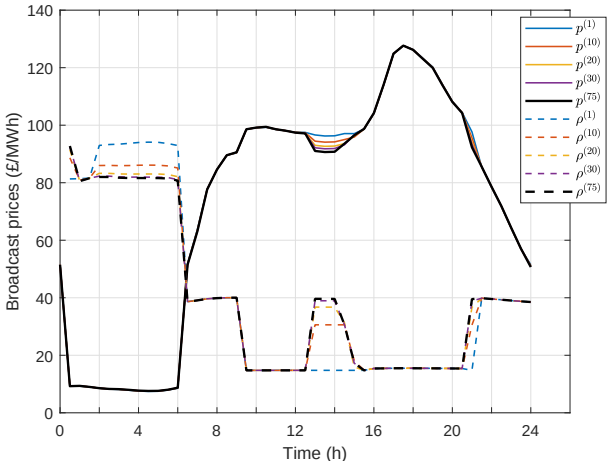
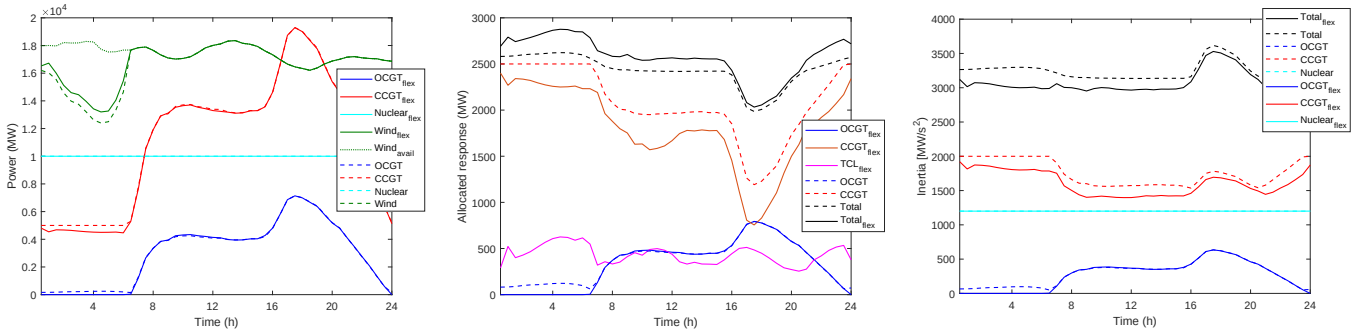


Fig. 6. Electricity price $p^{(n)}$ (bold lines) and response availability price $\rho^{(n)}$ (dashed lines) at different iterations n of the coordination algorithm.

impact on generation is marginal. However, the most noticeable changes occur during night hours, typically characterized by higher ancillary services requirements to account for low inelastic demand. Without TCL support, the optimal solution envisages a further curtailment of wind output (20.7% more, i.e. 4 GWh) in favor of an increase in OCGT and CCGT generation, as wind does not provide IR and FR. In absence of TCL response, it is necessary to commit and generate with OCGT, as IR and FR of cheaper CCGT have reached their maximum values. Note how the active participation of TCLs facilitates the integration of low-carbon energy, although this was not the explicit design objective of the proposed coordination scheme.

Results confirm the interplay between the allocation of IR and FR, with or without TCL flexibility [24]. In the latter case, the total FR decreases (dashed vs solid black line in Fig. 7b) while, in parallel, Fig. 7c shows a system-level increase of committed IR (dashed vs solid black line). Therefore, the lack of FR from TCLs is not generally compensated by the



(a) Generation profiles. (b) Breakdown of allocated frequency response. (c) Breakdown of inertial response.

Fig. 7. Relevant UC quantities when TCL flexibility is enabled (solid lines) and for a no-flexibility scenario (dashed lines).

same amount of FR from other technologies or, conversely, by more IR while maintaining constant FR from OCGT and CCGT. Instead, it is the result of a tradeoff between these two quantities.

Having discussed the aggregate impact of the TCLs population on the system through the quantities U_{TCL} and R_{TCL}, the behaviour of the single loads is now analysed. In particular, Fig. 8 shows the evolution across time of the internal temperature T of the individual TCL when the optimal power schedule u^* is applied, considering different values of initial temperature $T(0) = \tilde{T}$. It can be seen how, through the introduction of the auxiliary state variable I and associated cost terms in (8), temperature synchronicity is avoided. TCLs will in general have different internal temperatures, thus avoiding rebound peaks in the profiles of U_{TCL} and R_{TCL}, and facilitating convergence to a final stable configuration. Having chosen a terminal cost function Ψ that penalizes quadratically the quantity $I(t_{\mathrm{FIN}}) = T(t_{\mathrm{FIN}})) - T(0)$, the final temperature $T(t_{\mathrm{FIN}})$ of each TCL will tend to coincide with its initial value $T(0)$.

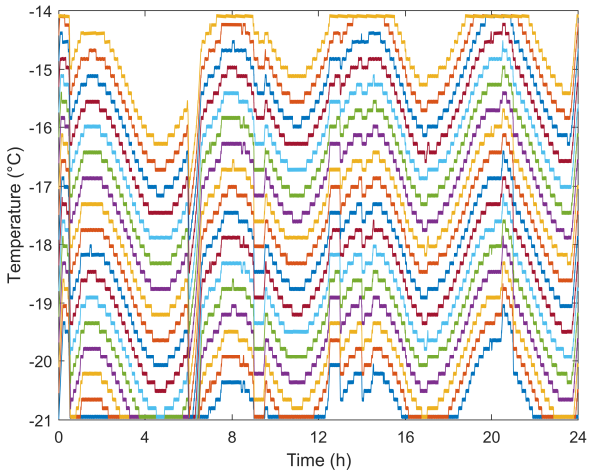


Fig. 8. Internal temperature T of individual TCLs (each with different initial temperature $\tilde{T}$) when the optimal power schedule u^* is applied.

Finally, the different cost components for the individual TCL at the equilibrium solution (as described in Section II-B) are represented in Fig. 9 as a function of the initial temperature $T(0) = \tilde{T}$. Note that the total daily cost (black) sustained by a single TCL is about 0.05£. As expected, the total cost will in general be lower for TCLs with lower initial temperature $\tilde{T}$. Since λ in (6) is monotone decreasing with respect to T, devices with lower temperatures in the first hours of the day (when FR prices are higher) will be able to allocate more FR and therefore receive higher revenues. It should also be emphasized that the impact of electricity costs and FR revenues in (8) is significantly larger than the tracking terms $\alpha(T(t) - \bar{T})^2$ and $\beta(I(t) - \bar{I}(t))^2$, which have been introduced to ensure coordination of the devices but do not significantly change the overall costs sustained by the TCLs. We recall that the presented results have been obtained with a temperature variation reference $\bar{I}(t) = \sin(t)$. The sensitivity of the results with respect to $\bar{I}(t)$ has been assessed: the variations in the price quantities represented in Fig. 9 are negligible (always below 0.002£) when different amplitude, magnitude and phase are considered for the sinusoidal reference.

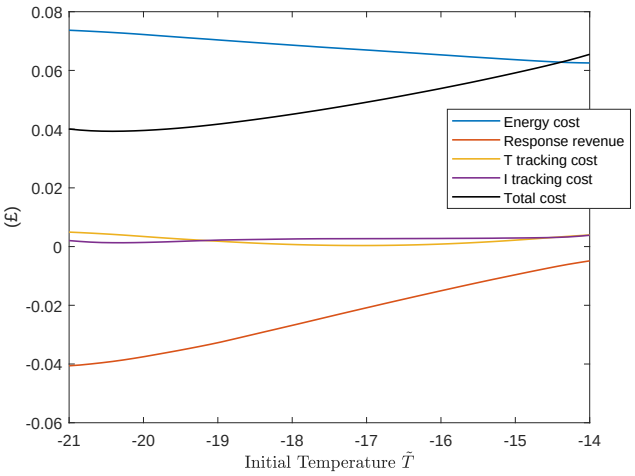


Fig. 9. Costs breakdown of single TCL as function of initial temperature $\tilde{T}$.

C. Performance Comparison

The results obtained with the proposed distributed competitive framework (hereby denoted as DCF) are now compared with two alternative scenarios:

1) *Business-as-usual* (BAU): the TCLs do not exploit their flexibility and they operate exclusively according to their

internal temperature T. They switch ON ($u(t) = P_{\mathrm{ON}}$) when they reach their maximum feasible temperature T_{MAX} and they switch back OFF again ($u(t) = 0$) when they reach the minimum temperature T_{MIN}.

2) *Cooperative-Centralized Framework* (CCF): the TCLs are flexible but they do not independently determine their power consumption profile. This is specified instead by a coordinating entity that centrally determines the behaviour of the whole population, with the global objective of optimizing TCLs participation and minimize the system operational cost. The cooperative coordination scheme considered in our comparison is the one presented in [8]. Please note that, in this paper, the term "decentralized" is used to characterize the implementation of the proposed control strategy, which remains cooperative and considers a centralized optimization.

The comparison between the three scenarios has initially considered the total system costs, which are represented in Fig. 10. As expected, a flexible use of the TCLs in the DCF and CCF scenarios can reduce the system generation costs (blue bars) by about 2% and 2.1%, respectively.

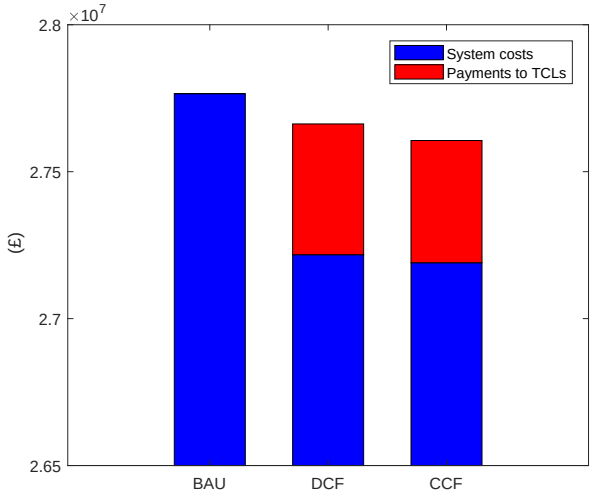


Fig. 10. Minimized system costs (blue) and payment to TCLs for FR allocation (red) in the three considered scenarios.

However, for a fair comparison between the three cases, one should also consider the payments that are made to the TCLs to reward them for their frequency response allocation (red bars). In both the DCF and CCF case, these simply correspond to $\int_0^{t_{\mathrm{FIN}}} \rho(t) R_{\mathrm{TCL}}(t)$, i.e. the integral over the considered time horizon of the total response R_{TCL} allocated by the TCLs, multiplied by the availability fee ρ. Note that, even when accounting for this additional term, the total system costs are still lower by 0.4% in the DCF and by 0.6% in the CCF when compared with the no-flexibility scenario BAU. The discrepancy in the cost savings corresponds to the so-called "price of anarchy" of the present set-up, i.e. the degradation of the whole-system performance when the individual agents compete between each other rather than cooperate towards the same high-level objective. We wish to emphasize that the prosed DCF strategy, albeit exhibiting a slightly higher system cost when compared to CCF, returns on the other hand a "fair" market equilibrium where each TCL is satisfied by the final outcome and cannot further reduce its own cost by modifying its initially planned strategy.

The three considered scenarios have also been compared in terms of daily costs sustained by the single TCLs, which are represented in Fig. 11 as a function of initial temperature $T(0) = \tilde{T}$. The calculation of the revenue from FR allocation, which constitutes a relevant (negative) component of costs, is different in the three cases. In the BAU scenario, the FR revenue is zero, since the inflexible TCLs do not provide frequency support. In the DCF scenario, it is sufficient to consider the relevant component $\int_0^{t_{\mathrm{FIN}}} \rho(t) r(T(t), u(t))\, dt$ in (8), whereas in the CCF case the FR revenue corresponds to $\frac{1}{N} \int_0^{t_{\mathrm{FIN}}} \rho(t) R_{\mathrm{TCL}}(t)\, dt$, i.e. an equal repartition among the N loads of the total revenues.

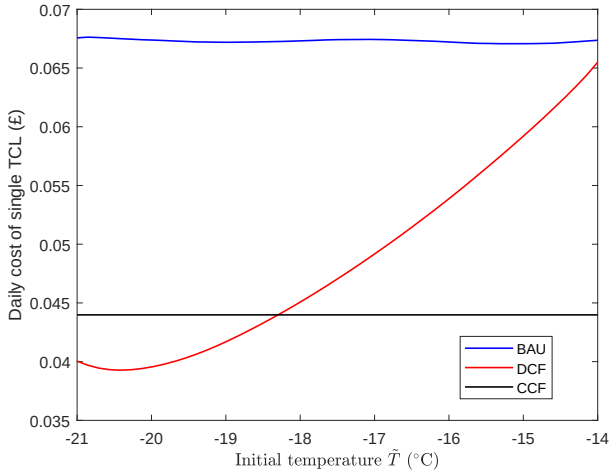


Fig. 11. Daily costs of individual TCLs, as a function of the initial temperature $\tilde{T}$, in the three considered scenarios.

It can be seen in Fig. 11 that, in the BAU case, the costs are more or less constant (since all loads are simply switching ON and OFF at regular intervals) and higher than in the DCF and CCF scenarios. In the CCF case, consistently with the considered cooperative framework, the TCLs costs have been equally shared among the loads and therefore they do not depend on the initial temperature $\tilde{T}$. Conversely, in the proposed competitive DCF case, the costs are impacted by $\tilde{T}$. As previously mentioned, a device with lower initial temperature will be able to consume more power during the early hours of the day, when electricity is cheap and there is significant value in frequency response provision. As a result, in the cooperative CCF solution, a certain group of TCLs (the ones with initial temperature below approximatively 18 °C) will see higher costs than in the DCF case and could, in principle, modify their schedule to achieve better results.

VII. Conclusions

In this paper, a novel approach for distributed price-based coordination of TCLs has been obtained by combining a rigorous mean field game formulation with a detailed modelling of system UC. As a result, a compact characterization of the

equilibrium solution for the TCLs coordination problem has been derived, enabling the provision of ancillary services by the TCLs and explicitly accounting for different generation technologies. Practical implementation methods and ad-hoc anti-synchronization schemes have been designed and tested in simulation, evaluating the potential benefits of the proposed control scheme for the single loads and for the overall system.

Future work will focus on a closer integration of TCLs' operation with other elements of the future smart home. In particular, the possibility of interacting with renewable generation (specifically PVs) will be investigated. The proposed analytical framework will be expanded to quantify the potential cost reductions obtained by TCLs that preferentially operate when "free" renewable energy from PVs is available. Interactions between TCLs and domestic storage will also be explored, with the purpose of a more effective energy arbitrage that makes use of storage capacity to overcome power consumption limitations of the TCLs (caused by their thermal constraints).

A more detailed modelling of system operation in comprehensive future energy scenarios will also be analysed. Finally, stochastic elements will be introduced in the model to account for exogenous and endogenous uncertainties (e.g. on prices and on TCL parameters) and to consider alternative control strategies, such as the use of random transitions by the TCLs between their ON and OFF state.

Acknowledgements

The authors would like to thank the anonymous reviewers for their helpful comments on the manuscript.

References

[1] J. L. Mathieu, M. Kamgarpour, J. Lygeros, G. Andersson, and D. S. Callaway, "Arbitraging intraday wholesale energy market prices with aggregations of thermostatic loads," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 763–772, March 2015.

[2] Y. Chen, A. Bui, and S. Meyn, "Estimation and control of quality of service in demand dispatch," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5348–5356, Sept 2018.

[3] P. Babahajiani, Q. Shafiee, and H. Bevrani, "Intelligent demand response contribution in frequency control of multi-area power systems," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1282–1291, March 2018.

[4] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1284–1293, Aug 2007.

[5] D. Angeli and P. Kountouriotis, "A stochastic approach to dynamic-demand refrigerator control," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 3, pp. 581–592, May 2012.

[6] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 189–198, Jan 2015.

[7] M. Song, C. Gao, M. Shahidehpour, Z. Li, J. Yang, and H. Yan, "State space modeling and control of aggregated tcls for regulation services in power grids," *IEEE Trans. Smart Grid*, pp. 1–1, 2018.

[8] S. H. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 5, pp. 1685–1700, Sept 2015.

[9] A. L. Motto, F. D. Galiana, A. J. Conejo, and M. Huneault, "On walrasian equilibrium for pool-based electricity markets," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 774–781, Aug 2002.

[10] K. Singh, N. P. Padhy, and J. Sharma, "Influence of price responsive demand shifting bidding on congestion and lmp in pool-based day-ahead electricity markets," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 886–896, May 2011.

[11] A. Khodaei, M. Shahidehpour, and S. Bahramirad, "Scuc with hourly demand response considering intertemporal load characteristics," *IEEE Trans. Smart Grid*, vol. 2, no. 3, pp. 564–571, Sept 2011.

[12] D. Papadaskalopoulos and G. Strbac, "Decentralized participation of flexible demand in electricity markets - part i: Market mechanism," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3658–3666, Nov 2013.

[13] ——, "Nonlinear and randomized pricing for distributed management of flexible loads," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1137–1146, March 2016.

[14] C. Gu, X. Yan, Z. Yan, and F. Li, "Dynamic pricing for responsive demand to increase distribution network efficiency," *Appl. Energy*, vol. 205, pp. 236 – 243, 2017.

[15] S. Li, W. Zhang, J. Lian, and K. Kalsi, "Market-based coordination of thermostatically controlled loads - part i: A mechanism design formulation," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1170–1178, March 2016.

[16] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 940–951, 2013.

[17] Z. Ma, D. Callaway, and I. Hiskens, "Decentralized charging control of large populations of plug-in electric vehicles," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 1, pp. 67–78, 2013.

[18] H. Chen, Y. Li, R. H. Y. Louie, and B. Vucetic, "Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1744–1754, July 2014.

[19] A. De Paola, D. Angeli, and G. Strbac, "Price-based schemes for distributed coordination of flexible demand in the electricity market," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 3104–3116, Nov 2017.

[20] Z. Zhu, S. Lambotharan, W. H. Chin, and Z. Fan, "A mean field game theoretic approach to electric vehicles charging," *IEEE Access*, vol. 4, pp. 3501–3510, 2016.

[21] A. De Paola, D. Angeli, and G. Strbac, "Distributed control of micro-storage devices with mean field games," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1119–1127, March 2016.

[22] D. Bauso, "Dynamic demand and mean-field games," *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6310–6323, Dec 2017.

[23] L. Zhang, T. Capuder, and P. Mancarella, "Unified unit commitment formulation and fast multi-service lp model for flexibility evaluation in sustainable power systems," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 658–671, April 2016.

[24] F. Teng, V. Trovato, and G. Strbac, "Stochastic scheduling with inertia-dependent fast frequency response requirements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1557–1566, March 2016.

[25] V. Trovato, A. Bialecki, and A. Dallagi, "Unit commitment with inertia-dependent and multi-speed allocation of frequency response services," *IEEE Transactions on Power Systems*, 2018 (in press).

[26] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2000.

[27] S. Wang, F. Gao, and L. Teo, "An upwind finite-difference method for the approximation of viscosity solutions to hamilton-jacobi-bellman equations," *IMA J. Math. Control Inform.*, vol. 17, no. 2, pp. 167–178, 2000.

[28] R. J. LeVeque, *Numerical Methods For Conservation Laws*, ser. Lectures in Mathematics. ETH Zurich. Birkhauser Basel, 1990.

[29] V. Trovato, S. H. Tindemans, and G. Strbac, "Leaky storage model for optimal multi-service allocation of thermostatic loads," *IET Gen. Transm. Dis.*, vol. 10, no. 3, pp. 585–593, 2016.

Antonio De Paola (M'15) received the M.Sc. degree in control engineering from the University of Rome Tor Vergata, Rome, Italy, in 2011, and the Ph.D. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2015. In 2017 he was awarded an Early Career Fellowship from the Leverhulme Trust and he is currently a Lecturer at the University of Bath. His main research interests include optimal control and game theory, with a specific focus on the design of novel control and market solutions for the deployment of a sustainable and decarbonized electricity grid.

Vincenzo Trovato (M' 15) is Research Engineer at the EDF Energy R&D U.K. Centre and a Visiting Researcher at Imperial College London, London, U.K. Previously he was Research Associate at Imperial College London where he obtained the Ph.D. degree in Electrical Engineering. He also owns a M.Sc. Degree in Electrical Engineering from Politecnico di Torino, Italy. His research interests include flexible assets and their impact on power system control and economics.

David Angeli (F'15) was born in Siena, Italy, in 1971. He received the B.S. degree in Computer Science Engineering and the Ph.D. in Control Theory from University of Florence, Italy, in 1996 and 2000, respectively. Since 2000 he was an Assistant and Associate Professor (2005) with the Department of Systems and Computer Science, University of Florence. In 2007 he was a visiting Professor with I.N.R.I.A de Rocquencourt, Paris, France, and since 2008, he joined as a Senior Lecturer the Department of Electrical and Electronic Engineering of Imperial College London, where he is currently a Reader in Nonlinear Control and the Director of the MSc in Control Systems. He is a Fellow of IEEE and of IET. He has served as an Associate Editor for IEEE Transactions in Automatic Control and Automatica. Overall he authored more than 100 journal papers in the areas of Stability of nonlinear systems, Control of constrained systems (MPC), Chemical Reaction Networks theory and Smart Grids.

Goran Strbac (M'95) is a Professor of energy systems at Imperial College London, London. His current research is focused on modelling and optimisation of economic and security performance of energy system operation and investment, including analysis of future energy markets to support cost-effective transition to smart low carbon energy future.